



Mellanox WinOF VPI Release Notes

Rev 4.60

Last Modified: 26 March, 2014

NOTE:

THIS HARDWARE, SOFTWARE OR TEST SUITE PRODUCT ("PRODUCT(S)") AND ITS RELATED DOCUMENTATION ARE PROVIDED BY MELLANOX TECHNOLOGIES "AS-IS" WITH ALL FAULTS OF ANY KIND AND SOLELY FOR THE PURPOSE OF AIDING THE CUSTOMER IN TESTING APPLICATIONS THAT USE THE PRODUCTS IN DESIGNATED SOLUTIONS. THE CUSTOMER'S MANUFACTURING TEST ENVIRONMENT HAS NOT MET THE STANDARDS SET BY MELLANOX TECHNOLOGIES TO FULLY QUALIFY THE PRODUCT(S) AND/OR THE SYSTEM USING IT. THEREFORE, MELLANOX TECHNOLOGIES CANNOT AND DOES NOT GUARANTEE OR WARRANT THAT THE PRODUCTS WILL OPERATE WITH THE HIGHEST QUALITY. ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NON-INFRINGEMENT ARE DISCLAIMED. IN NO EVENT SHALL MELLANOX BE LIABLE TO CUSTOMER OR ANY THIRD PARTIES FOR ANY DIRECT, INDIRECT, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES OF ANY KIND (INCLUDING, BUT NOT LIMITED TO, PAYMENT FOR PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY FROM THE USE OF THE PRODUCT(S) AND RELATED DOCUMENTATION EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.



Mellanox Technologies
350 Oakmead Parkway Suite 100
Sunnyvale, CA 94085
U.S.A.
www.mellanox.com
Tel: (408) 970-3400
Fax: (408) 970-3403

Mellanox Technologies, Ltd.
Beit Mellanox
PO Box 586 Yokneam 20692
Israel
www.mellanox.com
Tel: +972 (0)74 723 7200
Fax: +972 (0)4 959 3245

© Copyright 2014. Mellanox Technologies. All Rights Reserved.

Mellanox®, Mellanox logo, BridgeX®, ConnectX®, CORE-Direct®, InfiniBridge®, InfiniHost®, InfiniScale®, MLNX-OS®, PhyX®, SwitchX®, UFM®, Virtual Protocol Interconnect® and Voltaire® are registered trademarks of Mellanox Technologies, Ltd.

Connect-IB™, ExtendX™, FabricIT™, Mellanox Open Ethernet™, Mellanox Virtual Modular Switch™, MetroX™, MetroDX™, ScalableHPC™, Unbreakable-Link™ are trademarks of Mellanox Technologies, Ltd.

All other trademarks are property of their respective owners.

Table of Contents

Table of Contents	3
List of Tables	4
Chapter 1 Overview.....	5
1.1 Mellanox WinOF Rev 4.60 New Features.....	5
1.2 WinOF VPI Package Contents	5
1.3 Supported Operating Systems Versions	6
1.4 Supported Network Adapter Cards	6
1.4.1 Package Supplied Firmware	6
Chapter 2 Changes and Major New Features.....	7
2.1 New Features, Changes and Fixes in Version 4.60 From Version 4.55	7
2.2 New Features, Changes and Fixes in Version 4.55 From Version 4.40	9
2.3 New Features, Changes and Fixes in Version 4.40 From Version 4.3	10
2.4 New Features, Changes and Fixes in Version 4.3 From Version 4.2	12
2.5 New Features, Changes and Fixes in Version 4.2 From Version 3.2.0.....	13
Chapter 3 Beta Features	16
Chapter 4 Unsupported Functionality/Features.....	17
4.1 ConnectX®-2 Adapter Limitations	17
Chapter 5 Known Issues	18
Chapter 6 API Changes.....	29
6.1 API Changes in WinOF Rev 4.60	29
6.2 API Changes in WinOF Rev 4.55	29
6.3 API Changes in WinOF Rev 4.40	29

List of Tables

Table 1:	New Features, Changes and Fixes in v4.60	7
Table 2:	New Features, Changes and Fixes in v4.55	9
Table 3:	New Features, Changes and Fixes in v4.40	10
Table 4:	New Features, Changes and Fixes in v4.3	12
Table 5:	New Features, Changes and Fixes in v4.2	13
Table 6:	Beta Features	16
Table 7:	WinOF Known Issues	18
Table 8:	API Changes in WinOF Rev 4.60	29
Table 9:	API Changes in WinOF Rev 4.55	29
Table 10:	API Changes in WinOF Rev 4.40	29

1 Overview

This is the release notes for Mellanox WinOF Rev 4.60 VPI drivers.

The driver intends to improve performance and functionality of the Inbox driver. Therefore, Mellanox strongly recommends updating to its latest driver to achieve optimal performance and additional functionality.

1.1 Mellanox WinOF Rev 4.60 New Features

- Added support for the following:
 - SR-IOV Ethernet (beta level)¹
 - CIM installation as a standalone package
 - DSCP priority over IPv4¹
 - IRQ Dynamic moderation
 - Configuration changes saving upon Inbox and previous releases upgrade
 - New report for bad cables¹
 - iSCSI boot over IPoIB

1.2 WinOF VPI Package Contents

The Mellanox WinOF Rev 4.60 for Windows package contains the following components:

- Core and ULPs:
 - IB HCA low-level drivers (mlx4)
 - IB Access Layer (IBAL)
 - Ethernet driver (ETH)
 - IP over InfiniBand (IPoIB)
 - NetworkDirect (ND)
- Utilities:
 - OpenSM: InfiniBand Subnet Manager is provided as a sample code. The sample code is intended to allow users to test or bring-up the InfiniBand fabric without a management console / switch (to get started).
For cluster production environments, Mellanox's recommendation is to use a Managed Switch or the UFM-SDN Appliance.
 - Low level performance tools
- InfiniBand Diagnostics tools
- Software Development Kit (SDK)
- Documentation

1. Requires firmware v2.30.8000 and above.

1.3 Supported Operating Systems Versions

Supported Operating Systems and Service Packs:

- Windows Server 2008 R2 (64 bit only)
- Windows Server 2012 (64 bit only)
- Windows Server 2012 R2 (64 bit only)

1.4 Supported Network Adapter Cards

Mellanox WinOF Rev 4.60 supports the following Mellanox network adapter cards:

- ConnectX®-3 Pro and ConnectX®-3 Pro EN/ firmware v2.30.5000 and above
 - 10, 40 and 56 Gb/s InfiniBand (IB)
 - 10 and 40 Gb/s Ethernet
- ConnectX®-3 and ConnectX®-3 EN/ firmware v2.11.1250 and above
 - 10, 40 and 56 Gb/s InfiniBand (IB)
 - 10 and 40 Gb/s Ethernet
- ConnectX®-2 and ConnectX®-2 EN/ firmware v2.9.1200 and above
 - 10, 40 and 56 Gb/s InfiniBand (IB)
 - 10 and 40 Gb/s Ethernet



ConnectX® adapter cards are not supported by this release of WinOF.

1.4.1 Package Supplied Firmware

Mellanox WinOF Rev 4.60 provides the following firmware for Mellanox NICs:

- ConnectX®-3 Pro and ConnectX®-3 Pro EN¹
 - Firmware v2.30.8050
 - 40 and 56 Gb/s InfiniBand (IB)
 - 40 Gb/s Ethernet
 - Firmware v2.30.8000
 - 10 Gb/s InfiniBand (IB)
 - 10 Gb/s Ethernet
- ConnectX®-3 and ConnectX®-3 EN/ firmware v2.30.8000
- ConnectX®-2 and ConnectX®-2 EN/ firmware v2.9.1200

1. The previous package of WinOF 4.60 included ConnectX®-3 Pro firmware v2.30.8000.

2 Changes and Major New Features

2.1 New Features, Changes and Fixes in Version 4.60 From Version 4.55



This package version is 4.60.17738. The package contains mixed versions of drivers:

- Bus and eth version 4.60.17718.
- The mux driver version is 4.60.17729.
- The IPoIB driver version is 4.60.17736.

Table 1 - New Features, Changes and Fixes in v4.60

Category	Description
Installation/Upgrade	<ul style="list-style-type: none"> • Enabled configuration changes saving upon Inbox and previous releases upgrade • Enabled CIM installation as a standalone package • Fixed an issue occurred when uninstalling and reinstalling the driver. The ConnectX-3 Pro Ethernet device was displayed in the Device Manager with a yellow bang (!). • Fixed an issues enabling the package's execution in modify mode resulting in driver being disabled
Generic	<ul style="list-style-type: none"> • Fixed random parsing failures of string registry entries • Fixed compilation failure of "Hello_world" in the SDK • Fixed the return value of <code>ib_query_ca()</code> if failed to allocate resources for operation
Performance	<ul style="list-style-type: none"> • Added support to IPv6-to-all <code>nd_*_*</code> tests • Fixed CPU utilization report in <code>nd_*_*</code> tests • Fixed correct bandwidth peak results in <code>ibv_send_bw</code> with UD QP • Fixed sync problems of bidirectional mode in <code>ibv_read_bw/ibv_write_bw</code> • Fixed an issue reporting incorrect adapter type in performance tuning log file
RoCE	Fixed RoCE mode parsing
ND	<ul style="list-style-type: none"> • Added the ability to rearm a CQ in the kernel • Added the ability to handle LID changes • Changed connection timeout behavior. Added the <code>STATUS_CONNECTION_REFUSED</code> return value upon connection timeout. • Fixed missing completions when working with Completion Queue with single entry

Table 1 - New Features, Changes and Fixes in v4.60

Category	Description
IPoIB	<ul style="list-style-type: none"> Added the ability to handle LID changes Added support for iSCSI boot over IPoIB Fixed unexpected behavior upon QP asynchronous event Fixed bad completions of VMQ and NonVMQ modes in IPoIB Fixed a failure occurred when setting the IPoIB adapter value to "SA Query Timeout" Fixed propagation of the physical link disconnection to virtual (part_man) interface Performance improvements in latency Fixed using CQ after VMQ is closed Fixed bad completion of VMQ QP that was caused by malformed WR
Ethernet	<ul style="list-style-type: none"> Modified the CQ size to prevent CQ overrun Added DSCP support over IPv4 Added traffic profile Changed the report link speed zero in case of disconnected network adapter LBFO: <ul style="list-style-type: none"> Fixed the team's MAC address uniqueness in the subnet of the team in Windows 2008 R2 Fixed port channel teaming with CISCO switch and Fabric Extenders traffic loose in Windows Server 2008 R2 Fixed an issue related to packets sent with corrupted VLAN header when they were meant to be untagged Fixed unexpected behavior upon QP asynchronous event Stability fixes Performance improvements
WMI/CIM	<ul style="list-style-type: none"> Added ControlledBy association to IBPort Fixed ConformsToProfiles association for SoftwareIdentity and DriverIdentity Fixed execution of all tests which were running when executing Diagnostic tests on one instance Fixed a failure occurred when running MLNX_Card Fixed the printing of diagnostics log Fixed an issue preventing from get-event to show information after disabling the PCI device Removed support for the following configuration: <ul style="list-style-type: none"> ModeFlags SingleMsixNum MultiMsixNum SingleEqNum MultiEqNum MaxContQuant SlaveNum DebugLevel DebugFlags UsePrio NumFcExch EnableQoS BlockMcastLoopBack InterruptFromFirstPacket ProbeVf

2.2 New Features, Changes and Fixes in Version 4.55 From Version 4.40

Table 2 - New Features, Changes and Fixes in v4.55

Category	Description
Generic	<ul style="list-style-type: none"> Added support for Windows Server 2012 R2 Operating System Added the <code>ParentBusPath</code> option to each port registry key Added a new hardware ID for ConnectX®-3 Pro NICs The QP numbers allocation is now round-robin manner <code>RecvCompletionMethod</code> as Interrupt is no longer supported Removed the <code>LsoV1IPv4</code> from the registry/UI Removed from the bus driver configuration the '<code>Non-DMA</code>' option Removed the <code>TXRingNum</code> option from the UI
NVGRE	<ul style="list-style-type: none"> Added NVGRE hardware off-load support Added to the UI the <code>*EncapsulatedPacketTaskOffload</code> option when using ConnectX®-3 Pro NICs
Performance	<ul style="list-style-type: none"> Added the <code>nd_send_bw</code> and <code>nd_send_lat</code> ND benchmarking tools Fixed <code>nd_*_bw</code> to achieve better performance (memory buffer alignment) and consistent results
Ethernet	<ul style="list-style-type: none"> Fixed the issue preventing messages to be sent in VLAN 0 when using many VMQ rings Added IP-IP checksum off-load support Added Ports TX arbitration/Bandwidth allocation per port The following ND providers, <code>MLX4ND</code> and <code>MLX4ND2</code> are installed by default
InfiniBand	<ul style="list-style-type: none"> IPoIB performance improvements Fixed a <code>part_man</code> issue related to wrong statistics over virtual partman interfaces
RoCE	<ul style="list-style-type: none"> Added Sniffer for RoCE packets The used RoCE mode set upon driver load is printed into event log message

2.3 New Features, Changes and Fixes in Version 4.40 From Version 4.3

Table 3 - New Features, Changes and Fixes in v4.40 (Sheet 1 of 3)

Category	Description
General	<ul style="list-style-type: none"> Added a notification in the event log in case SMB is not supported in ConnectX®-2 firmware Added the trace tool for WPP tracing Added copyright to the SDK files Added WMI/Powershell support Fixed an issue causing the setup to fail upon <code>perf_tuning</code> failure during the installation. An error message will be printed in the installation log upon <code>perf_tuning</code> failure. Removed port setting registry key during uninstall Fixed an issue with the Mellanox adapter being shown on the USB removal menu, which caused the removal of the Mellanox adapter once removing the USB.
Performance	<ul style="list-style-type: none"> Set 512 RX buffers by default Removed <code>TXRingNum</code> Changed the <code>perf_tuning</code> setting to achieve a better performance tuning Added the <code>nd_write_bw/nd_write_lat</code> and <code>nd_read_bw/nd_read_lat</code> tools Fixed the <code>perf_tuning</code> indication of the last chosen tuning scenarios Fixed a crash in the <code>ib_send_lat/bw</code> utilities caused when the port link was down Fixed the “Restore to defaults” option in the <code>perf_tuning</code> tool. Now the default values are being restored

Table 3 - New Features, Changes and Fixes in v4.40 (Sheet 2 of 3)

Category	Description
Ethernet	<ul style="list-style-type: none"> • Added Transmit Side Scaling (TSS) • Added Ethernet QoS proprietary counters, diagnostics and traffic for monitoring, using Windows' perfmon utility • Added to the MTU size the IP header size (1500 ->1514, 9600->9614). Thus the minimum Jumbo frame size is 614. • Interrupt moderation supports the following profiles: <ul style="list-style-type: none"> • Low Latency • Moderate • Aggressive <p>In addition to old values that are not supported anymore.</p> • Made mlx4_bus and Ethernet devices removable • Network Direct: Added support for NDv2 • Network Direct: Set the default ND provide value to mlx4nd2 • Fixed WoL support on NIC with a single port • Fixed the default RoCE configuration on NICs with a single ports • Fixed the values for the MTU and rate of the CM-REQ • Fixed miniport reset on sending scenarios • Removed the QoS attributes when disabling QoS • Enabled MaxRssProcessors support of the following values: 1, 2, 4, 8, 16, 32, 64 • Network Direct: Fixed a crash occurred when more than 4 SGEs elements were used in an ND write operation • Network Direct: Fixed the swap of InboundReadLimit and OutboundReadLimit when creating an EndPoint and in Connector::GetConnectionData • Network Direct: Fixed disallowing creation of EndPoint with zero attributes in the Receive Queue • Network Direct: Removed the option of NDK registration failure requiring a reboot of the machine to register it again • Network Direct: Fixed a failure when creating an EndPoint with zero attributes in the Receive Queue • Network Direct: Added the option of sensing the incoming Read messages according to the device capabilities when creating an EndPoint limit • Network Direct: Fixed a failure of ND connectivity between VMs on the same host

Table 3 - New Features, Changes and Fixes in v4.40 (Sheet 3 of 3)

Category	Description
InfiniBand	<ul style="list-style-type: none"> On rare occasions, depends on the GUID assignment, the IPoIB MAC address can be assigned with a multicast MAC (the least significant bit of the most significant address octet is set to 1). In that case, all of the traffic over the IPoIB I/F is dropped. If you experience this issue, please contact Mellanox support. Added <code>active_mtu</code> field to struct <code>ib_port_attr_t</code> Added the option of <code>vstat</code> displaying the <code>active_mtu</code> of the ports Allowed registration of a large Memory Region which is splitted to many segments Fixed a bluescreen issue that occurred when disabling the interface after a TX stress over the VMQ Fixed a failure of MPI/ND over InfiniBand Added the option of <code>ibv_devinfo</code> displaying the correct MTU value after it was changed Added the option of <code>part_man</code> printing the adapter name when the Port GUID is set to zero. Added the option of <code>part_man</code> printing the leading zeroes of port GUID
Installation/Upgrade	<ul style="list-style-type: none"> Prevented displaying a message to upgrade the firmware for OEM NICs if it has the latest firmware version Removed portsetting registry key during uninstall

2.4 New Features, Changes and Fixes in Version 4.3 From Version 4.2



WinOF VPI version 4.3 was released as an intermediate release.

Table 4 - New Features, Changes and Fixes in v4.3

Category	Description
General	<ul style="list-style-type: none"> Added support for a new provider called MLX4ND, which supports both NDv1 and NDv2 interfaces
Performance	<ul style="list-style-type: none"> Enabled performance tuning running according to the operating systems that are running over it. The keywords added to the registry in NDIS support Windows 2012 are: <ul style="list-style-type: none"> RssMaxProcNumber NumRSSQueues RSSProfile The rest of the keywords are added in all versions of NDIS. This change is based on: http://msdn.microsoft.com/en-us/library/windows/hardware/ff570864(v=vs.85).aspx

Table 4 - New Features, Changes and Fixes in v4.3

Category	Description
Ethernet	<ul style="list-style-type: none"> RoCE MTU value is no longer set to 1024 by default. All options stay as they are and can only be chosen if they were selected explicitly in the UI/registry. The current default state is as follows: The value is now derived from the MTU (or MaxFramSize, or Jumbo Packets value) and they are all aliases for the same value). The value is aligned to 256,512,1024,2048 in a way that it will be rounded down to the nearest power of two of the ETH MTU.
InfiniBand	<ul style="list-style-type: none"> Added ibdiagnet utility support

2.5 New Features, Changes and Fixes in Version 4.2 From Version 3.2.0

Table 5 - New Features, Changes and Fixes in v4.2 (Sheet 1 of 3)

Category	Description
General	<ul style="list-style-type: none"> Modified RSS cores and changed VMQ affinity on the fly Fixed restart issue when there are not enough MSI-X vectors for each machine core Added support for K-GROUPS processors (more than 64 processors support) to allow assignment of MSI-X affinity for multiple processor groups. Set an adequate number of MTTs to map all physical memory Allocated firmware and ICM memory in chunks of non-paged memory instead of using contiguous physical memory. Fixed RSS indirection table mapping building when there are less RX rings than RSS cores. Fixed a bug, preventing standard work with BAR value more than 4GB. Fixed memory leaks Fixed error flows causing a Bluescreen in driver startup/unload Fixed a Bluescreen occurrence upon shutdown due to leak in active resources Changed device names in device manager and their hardware IDs. The changes were made to distinguish between ConnectX®-2 and ConnectX®-3: <ul style="list-style-type: none"> for ConnectX-2: MLX4\ConnectX-2_Eth and IBA\ConnectX-2_IPoIB for ConnectX-3: MLX4\ConnectX-3_Eth and IBA\ConnectX-3_IPoIB Set QoS settings only for ConnectX-3. Changing the hardware ID, forces the OS to install new device and re-build the registry keys. Added an event log to indicate driver failure upon start if there are two HCA burned with the same GUID. Added firmware upgrade support as part of the setup process. The setup burns the new firmware only on Mellanox cards. Firmware burning failure does not prevent the driver's installation, therefore, it will show a warning. In this case, it is recommended to update the firmware manually. Enabled configuration of TxRingNum registry key from the UI Improved the "Port Protocol" dialog Added Registry key documentation to the setup package

Table 5 - New Features, Changes and Fixes in v4.2 (Sheet 2 of 3)

Category	Description
Performance	<ul style="list-style-type: none"> • Optimized code performance • Increased send parallelism • Memory used in receive flow is now allocated with the same affinity of the handling processor for faster access • Statistics parameters are now directly read from hardware instead of being calculated by software. • Added support for BlueFlame. BlueFlame is now the default working mode for all packets that have a descriptor which fits into a BF register (currently 256 bytes). Use "BlueFlame" registry key to enable/disable this feature. • Added support for RSS functionality on available processors numbers. Used to be restricted to start at the first processor. • Changed RSS registry defaults to give better out of the box performance • Added a performance UI to tune performance under various scenarios • Added a tool to tune performance under various scenarios
Ethernet	<ul style="list-style-type: none"> • Added support for multiple TX rings • Added an option to verify that the number of multicast groups used is no higher than the firmware limits • Improved performance in virtualization when using VMQ

Table 5 - New Features, Changes and Fixes in v4.2 (Sheet 3 of 3)

Category	Description
IPoIB	<ul style="list-style-type: none"> • Fixed a bug that prevented IBAL applications from working on machines with 2 cards • Fixed a bug that caused packet drop on remote node when the first sent packet is not an ARP or a multicast packet • Added support for multiple TX rings. The new driver advertises its TSS capability (using multiple TX rings). This feature will not be used when communicating with older version drivers (such as the inbox driver). This way IPv4 communication in a mixed environment that includes the inbox driver should work. IPv6 communication may be problematic in such an environment. • Eliminated some race conditions in IPoIB mcast handling • Added mcast addresses according to the Windows mcast table in addition to mcast addresses from igmp/mld queries, in order to connect to such mcasts that are not reported by the queries. • Removed redundant LID_CHANGE events. This fixes the redundant reconnection of all IPoIB mcasts. • Added support for VM migration. When a VM is migrated from one host to another it does not send a new ARP to hosts it already knew when it was sitting on the first host. In order to prevent communication interruption to such hosts, the driver will initiate an ARP to learn them again. • Fixed IPoIB VMQ affinity update bugs • Fixed IPoIB VMQ parent queue management race during reset • Fixed a bug in passing DHCP packets from a Linux VM • Fixed bug in RDMA statistics for NDK • Removed affinity restriction to group 0 for NDK • Added support for the part_man tool. It can create only one additional IPoIB interface per port GUID. This feature can be utilized to use different IPoIB interfaces for SMB and Hyper-V. The virtual ports configuration is removed during uninstall • Installed ND by default • The following tools were added to the setup package: <ul style="list-style-type: none"> • ibv_asyncwatch.exe • ibv_atomic_bw.exe • ibv_atomic_lat.exe • ibv_devinfo.exe • ibv_rc_pingpong.exe • ibv_rdma_bw.exe • ibv_rdma_lat.exe • ibaddr.exe • ibcacheedit.exe • iblinkinfo.exe • ibqueryerrors.exe • ibsysstat.exe • saquery.exe • smpdump.exe

3 Beta Features

Table 6 - Beta Features

Category	Description
WinVerbs	WinVerbs is currently at beta level.
ibdump	ibdump is currently at beta level.
SR-IOV Ethernet	SR-IOV Ethernet is currently at beta level.

4 Unsupported Functionality/Features

The following are the unsupported functionality/features in WinOF Rev 4.60:

- ND over WinVerbs provider
- SRP

4.1 ConnectX®-2 Adapter Limitations

- Limited NDK support prevents SMB-Direct support- for details contact Mellanox Support.
- CQ to EQ feature is not supported - consequently, the driver may associate RSS and VMQ packets with the wrong CPUs
- Mellanox perfmon counters are not supported
- Auto-sensing is not supported
- ibdump is not supported
- QoS is not supported

5 Known Issues

Table 7 - WinOF Known Issues (Sheet 1 of 11)

Category	Description	Workaround
General	Pinning all the physical memory (used by RDMA operations, such as register memory, pin user memory) on the machine, on Operating Systems prior to Windows Server 2012, may cause the machine to hang.	Avoid pinning the whole machine memory in those Operating Systems.
	When the tracer tool traces many events, it may consume a large amount of memory (up to several GB ram).	Reduce the verbose level.
	The displayed MAC address in the DHCP server of virtual IPoIB I/F may display a wrong data (FF00.....) although the I/F is still fully functional.	-
	When running applications that use ND or libibmad (such as OpenSM) the system might get to a unstable state when trying to shutdown/restart/hibernate it.	Close all applications that uses ND or libibmad before performing shutdown/restart/hibernate.
	Activating NC-SI in WinOF v4.60 may cause driver's loading failure when using an older firmware version than 2.30.8000.	Do not enable NC-SI in machines that WinOF v4.60 is installed in.
	When trying to use the following WMI classes and operations, an error or empty messages will be displayed: <ul style="list-style-type: none"> MLNX_Card MLNX_FirmwareIdentity MLNX_FirmwareInstallationService MLNX_Realizes 	-
	The maximum values returned by the <code>ib_query_ca()</code> function (for example: <code>max_qp</code> , <code>max_mr</code>) are the upper limits of the supported resources by the device. However, it may be impossible to use these maximum values, since the actual number of any resource that can be created may be limited by the machine configuration, the amount of host memory, user permissions, and the amount of resources already in use by other users/processes.	-
	Running Ntttcp without the "-a X" flag (X>1) in a NIC configured with 10GbE, may cause low bandwidth in TCP single stream.	Run Ntttcp with "-a 8" for best performance
	Active links disappear after changing the cable connectivity from Ethernet to InfiniBand or vice versa.	Disable and enable the <code>mlx4_bus</code> interface from the Device Manager.

Table 7 - WinOF Known Issues (Sheet 2 of 11)

Category	Description	Workaround
	On ConnectX®-2/ConnectX®-3 Ethernet adapter cards, there is a mismatch between the GUID value returned by firmware management tools and that returned by fabric/driver utilities that read the GUID via device firmware (e.g., using ibstat). Mlxburn/flint return 0xffff as GUID while the utilities return a value derived from the MAC address. For all driver/firmware/software purposes, the latter value should be used.	Please use the GUID value returned by the fabric/driver utilities (not 0xffff).
	When working with the default NDv1 and NDv2 providers, the following error message might be displayed: 0xC0000120 NT_STATUS_CANCELLED This error does not affect any functionality and can be safely ignored.	-
	Changing the default ND providers may cause random errors, such as: 0xC0000238 (NT_STATUS_ADDRESS_ALREADY_ASSOCIATED) on Connect() or with 0xC0000236 (NT_STATUS_CONNECTION_REFUSED) on Accept(). These errors can be safely ignored.	-
	The UI displays the network interface's state incorrectly even after disabling and re-enabling the "disconnected" network interfaces.	To see the correct state and link speed, perform one of the following: <ul style="list-style-type: none"> • Run Get-netadapter Powershell cmdlet or • Right click on that interface from "Network Connections" and click on status
	Upon bad Cable Recoverable the driver will print in the event log "Unsupported cable message" and upon unsupported cable the message be "Bad Cable".	-
	winctxtune.dll, the user interface module of the driver, is not digitally signed	-
	The winctxtune.dll version may be different from the driver's version (4.60.17718 or 4.60.17736)	-

Table 7 - WinOF Known Issues (Sheet 3 of 11)

Category	Description	Workaround
InfiniBand	The installation process does not close any applications running in the background, and may cause a BSOD as a result of a stuck cmd.	It is recommended to close all running applications prior to upgrading the driver.
	InfiniBand application that was compiled with an SDK version earlier than WinOF v4.40 is not binary compatible.	Recompile InfiniBand application with WinOF v4.40 and above. ND application is backward compatible and older applications over ND do not have to be recompiled.
	WMI does not work due to lack of permissions.	Change the execution policy. Run: <code>Set-ExecutionPolicy AllSigned</code>
	WinOF 4.40 and above IPoIB driver is not IPv6 compatible with earlier driver versions, including Window 8 Inbox driver. If WinOF 4.50 IPoIB node receives an icmpv6 message from the legacy IPoIB node, the following event will appear in the event log: "IPoIB driver detected a possible incompatibility with Windows 8 inbox IPv6 support due to which there is no IPv6 connectivity".	To enable compatibility mode, add: <ul style="list-style-type: none"> Win8InboxCompatibilityMode REG_SZ registry key with the value of 1 in the IPoIB interface registry. Note: All IPoIB nodes must use the same mode to enable IPv6 compatibility with earlier driver versions. We recommend upgrading all hosts to the new driver version from http://mellanox.com or use Windows 8 compatibility mode
	Functionality such as CQ to EQ; Auto Sensing, QoS are not supported in ConnectX®-2 adapter cards.	-
	On rare occasions, as a result of port configuration change (IB/ETH) the UI may get stuck for up to a few minutes. This effect does not require any user action. The UI returns to its proper functionality after a few minutes.	-
	WinOF Inbox driver does not support upgrade. When installing WinOF v4.40 and above on a Windows Server 2012 and above machine, the Inbox driver is uninstalled prior to starting the new installation and any previous configurations is lost. The Inbox driver will be reinstalled automatically when the new driver is uninstalled.	-
	Without separate ports for each stream, WinSock multiplexes every packet to every subscriber socket and then filters it out.	Use different UDP ports to get higher performance when using multicast packets.

Table 7 - WinOF Known Issues (Sheet 4 of 11)

Category	Description	Workaround
	A virtual IPoIB interface, created by the <code>part_man</code> utility, reports an Active state when the physical link is in the Initializing state and OpenSM is not running in the subnet	-
	The "Packets Received Discarded" and "Packets Received Errors" counter may display wrong results.	-
	Firmware upgrade may fail during installation if there was a prior firmware upgrade on the machine without a reboot after it. A firmware upgrade failure does not fail the whole installation.	Upgrade the firmware manually.
	The drivers' functionality is limited up to 1TB of memory and up to 128 cores.	-
	Connection failure on ND tests while machine A have IBAL provider and machine B have MLX4ND provider.	-
	Hibernate and Sleep are not functional when user-space is using its resources.	-
	Calling <code>ib_join_mcast()</code> with <code>timeout_ms = 0</code> may result in a BSOD.	Avoid calling this function with <code>timeout_ms = 0</code> .
	IPoIB does not support: <ul style="list-style-type: none"> • MAC address change • QoS (packet priority) • Load balancing and fail-over (LBFO) • Connected Mode • Partition 	-
	Memory registration on 32bit machines is limited to up to 256GB.	-
	In an interoperability environment that has both Linux and Windows OSs, the MTU value must be the same, otherwise packets larger than the minimum will not go through. The default MTU for Linux is 2K and for Windows is 4K.	-
	Old InfiniBand application, which was complied with WinOF earlier than 4.40, must be recompiled with the latest SDK version.	-

Table 7 - WinOF Known Issues (Sheet 5 of 11)

Category	Description	Workaround
	OpenSM does not run as a service during installation since the SM requires the GUID parameter to decide which port to work on. Setting it on setup causes it to work only on the first port and not the others.	To run OpenSM as a service, assuming the package was installed in the default path, use: <pre>sc create opensm bin-Path= "c:\Program Files\Mellanox\MLNX_VPI\IB\Tools\opensm.exe"</pre> To start the service run: <pre>sc start opensm</pre>
	Tools issues: <ul style="list-style-type: none"> ibportstate does not work on RoCE ports ibdiagpath may crash on Hyper-V machines 	-
	If an application which uses InfiniBand runs while the driver is being restarted, a bluescreen or an NMI may occur.	Stop all InfiniBand applications including OpenSM upon driver restart.
	If OpenSM is up during driver restart on the same machine, it might stop working when the driver is being backed up.	To resume operation, stop OpenSM and restart the driver while OpenSM is down.
	When configuring one of the ports to be IPoIB, Mellanox device appears in the Eject Devices.	-
	Sometimes during traffic, the latency of the IPoIB can get higher (this behavior is inconsistent).	Set the following registry in the IPoIB network interface: RecvIntModTime=0 Please note, this action increases the CPU utilization.
	No communication between the physical interface and a VM that uses vSwitch created over virtual IPoIB (Pkey), and vice versa.	-
Ethernet	Disabling the "Priority & VLAN tag" in the UI which VLANID is configured, may result in sending packets with the configured VLANID.	Remove the VLANID before disabling the "Priority & VLAN tag".
	When working with LBFO, the teamed interface disappears after machine reboot. The issue applies to Windows Server 2012 R2	Delete the existing teamed interface and create a new one.
	WakeOnMagicPacket registry key is not added to the registry although WoL is supported by the driver and by the NIC.	-
	When the ports of the device are configured as Ethernet only, ibstat/vstat may display wrong information.	-

Table 7 - WinOF Known Issues (Sheet 6 of 11)

Category	Description	Workaround
	High multicast drop rate on multicast storming.	Use "Multicast traffic" tuning option under the performance tab. For further information, please refer to section "Tunable Performance Parameters" in the User Manual.
	Driver installation requires deletion of the <code>mlx4_bus.sys</code> file in Windows Server 2008 R2 and WLH OSs when using the PXE package.	Delete the <code>mlx4_bus.sys</code> file and reboot the machine to install the driver.
	When there is a stress in TCP connection establishments, some of those connections may fail.	Increase the Ring queue sizes: <ul style="list-style-type: none"> • ReceiveBuffers - controls the receive ring size • TransmitBuffers - controls the transmit ring size
	Wake on Lan (WoL) cannot be disabled on NICs which supports it.	-
	The DCB component specifies a default traffic classification that is applied to all egress packets that do not match other classification conditions. In this case, the network adapter assigns the IEEE 802.1p priority level that is associated with the default classification to these egress packets. The default traffic classification has the following attributes: <ul style="list-style-type: none"> • It has a traffic classification condition of type <code>NDIS_QOS_CONDITION_DEFAULT</code>. • It is the first traffic classification defined in the array of <code>NDIS_QOS_CLASSIFICATION_ELEMENT</code> structures. 	-
	RDMA Activity counters do not count during Network-Direct RoCE traffic.	
NVGRE	When NVGRE off-load is enabled, the GRE traffic cannot be accepted as a regular L2 traffic and requires special <code>L2_TUNNELING</code> steering rules. In such case the GRE packets are dropped or directed to promiscuous queue.	-
	GRE traffic steering by inner MAC and by outer MAC simultaneously is currently not supported.	Configure steering or by inner MAC, or by outer MAC.

Table 7 - WinOF Known Issues (Sheet 7 of 11)

Category	Description	Workaround
	If VMQ set filter requests are accepted without a GRE flag (i.e. requested steering by outer MAC), the GRE packets do not reach that VMQ.	Set the bus driver registry keyAcceptGREbyOuterMAC_P1/2 per port to accept GRE traffic by outer MAC and to duplicate L2 steering rule to L2_TUNNELING rule for each VMQ set filter request without GRE flag. Note: For regular NVGRE Hyper-V scenarios the value of the registry key below must be set to 0:AcceptGREbyOuterMAC_P1/2
Quality of Service	Running Quality of Service (QoS) commands without the parameter "-PolicyStore ActiveStore" may cause machines to load without Quality of Service policy.	Store the QoS policy in the ActiveStore
RoCE	RoCE does not support: <ul style="list-style-type: none"> Traffic cannot go through the router. It works in the same subnet only Multicast traffic VLAN Layer 3 feature 	-
	Using different versions of RoCE in your cluster is not supported.	Use the same RoCE version in all the cluster in the Ethernet ports.
Performance	When using WinOF 4.40 or above, low throughput will be seen on 40GbE adapters when QoS is enabled.	Disable QoS when it is not in use: <ul style="list-style-type: none"> Open a PowerShell prompt. Run: Disable-NetAdapter-Qos -name <Interface Name> where <Interface Name> is e.g. "Ethernet 1"
	perf_tuning is supported only when one of the two NUMA nodes are in use.	-
	Execution of nd_write_lat over mlx4ndv1 or mlx4ndv2 may never end.	Press CTRL-C and kill the test.

Table 7 - WinOF Known Issues (Sheet 8 of 11)

Category	Description	Workaround
	Running performance benchmarks for a short period of time (< 1 sec) may provide bad latency in IPoIB and Ethernet.	Set "Rx Interrupt Moderation Profile" and "Tx Interrupt Moderation Profile", to "Low Latency" to avoid bad latency. Note: This may increase CPU utilization.
Hyper-v	When the vSwitch is detached from the ETH\IPoIB device while the driver is disabled, the device does not reacquire the static IP it had before the attachment of the vSwitch. When the vSwitch is attached to the ETH\IPoIB device while there is no link, it will not receive the device IP when the link is back up.	-
	After attaching the vSwitch to the ETH\IPoIB device, changing the "Jumbo Packet" registry key on the ETH\IPoIB device does not affect the vSwitch configuration and vice versa. For example, if the user sets the "Jumbo Packet" on the ETH\IPoIB device to X, and the "Jumbo Packet" on the vSwitch to X+Y, X+Y sized packets will be passed from NDIS down to the driver and they will be dropped by it.	Reattach the vSwitch to sync with the value set in the ETH\IPoIB device.
	Unexpected behavior might occur when running in a virtualized environment and creating two virtual switches bound to each of the ports of a dual port NIC and then using both of them with two vNICs from the same VM.	-
	When moving an IPoIB interface in a VM from non-VMQ to VMQ or from VMQ to non-VMQ, a reset to the NIC may occur and in the event log the following message will appear: 'device reports a "CQE error" on cqn <number> qpqn <number> Status <number>. Therefore, the HCA Nic will be reset. (The issue is reported in Function <function>). For more information refer to details.'	-
	In IPoIB when using long Multicast traffic from a Virtual Machine (VM) to an external host there might be up to 0.5% loss in 5% bursts	-
	Hyper-V is at low bandwidth on LBFO vSwitch, Windows 2012	
	In Ethernet to achieve better iperf TCP performance between a Linux VM and a Windows VM on different hosts, when using MS MUX over the Ethernet driver, use the non VMQ mode for the VMs.	-

Table 7 - WinOF Known Issues (Sheet 9 of 11)

Category	Description	Workaround
	After disabling and enabling a port on a guest, a ping to it may be renewed after a minute. The ARP requests sent by Windows are less frequent as the time passes. If the guest port was down for a while, it could take time until Windows decides to send another ARP request to it.	-
	When VMQ is enabled after reset, the driver loads all the VMQs that existed before the reset. However, it is not guaranteed that each VMQ will receive the same QP number it had before the reset. This can cause some delay as a result of resetting before connectivity is reestablished. The delay is caused by the time it takes for the ARP table to update after initiating the Gratuitous ARP.	-
	The IPoIB non-VMQ mode is supported only when the VMQ is enabled according to the registry values.	To use the non-VMQ mode for a VM, change its settings as follow: <ul style="list-style-type: none"> • Press "Settings" on the VM • Go to Network Adapter -> Hardware Acceleration • Un-check the "Enable virtual machine queue"
Installation/ Upgrade	Upgrading the driver while the UI is opened with the "ConnectX NIC device" may cause the installation process to never end.	Close the UI before driver upgrade.
	Rebooting the machine reboot while uninstalling WinOF, may result in installation failure.	Delete Mellanox components from HKEY_LOCAL_MACHINE\SOFTWARE\Microsoft\Windows\CurrentVersion\DIFxApp\Components. The Mellanox components are mlx4eth63, ipoib6x and mlx4_bus
	Canceling the installation process may leave the bus driver in a disable state. The driver appears in a yellow bang containing the following error message: "Windows cannot start this hardware device because its configuration information (in the registry) is incomplete or damaged. (Code 19)".	Scan for new hardware and reboot the machine.
	Downgrade is not supported.	Uninstall the current version and install the older one.

Table 7 - WinOF Known Issues (Sheet 10 of 11)

Category	Description	Workaround
	Installation is stuck when Remote Desktop Session Host Windows Installer RDS Compatibility is enabled.	Disable the Remote Desktop Session Host Windows Installer for the duration of the installation according to the procedure described in: http://support.microsoft.com/kb/2655192/
	Occasionally, upon driver upgrade due to mismatch of driver versions, the following error message is displayed: "There is an interface mismatch between ETH driver and the bus driver. The ETH driver interface version is 50 while the bus driver interface version is 17. As a result the ETH driver has failed to start. This happened due to setup failures or partial update of the drivers. In order to resolve the issue, please reboot the computer" The message can be safely ignored.	-
	Configuration is not restored when replacing a ConnectX-2/ConnectX-3 NIC with a ConnectX-3 Pro NIC located on the same PCI slot	Clean the old network adapter configuration prior to upgrade.
	Configuration can be restored only in Windows Server 2012 and above	-
	IPv6 configuration restore is not supported	-
	Upon upgrade, the following Registry Key values will be overwritten with the following: <ul style="list-style-type: none"> • *ReceiveBuffers = 512 • *MaxRssProcessors = 8 • *RssBaseProcNumber = 0 • *NumRSSQueues = 8 • *RssMaxProcNumber = 63 • *RssProfile = 1 • DefaultRecvRingProcessor = -1 • TxInterruptProcessor = -1 • TxForwardingProcessor = -1 • RxIntModerationProfile = 1 • TxIntModerationProfile = 1 • RecvCompletionMethod = 1 • SingleStream = 0 • TxRingNum = 8 	-

Table 7 - WinOF Known Issues (Sheet 11 of 11)

Category	Description	Workaround
	Upon upgrade the following Ethernet Registry Keys will be deleted: <ul style="list-style-type: none"> • SendCompletionMethod • UseRSSForRawIP • UseRSSForUDP 	-
	Upon upgrade the SendCompletionMethod IPoIB Registry Key value will be modified as follow: <ul style="list-style-type: none"> • SendCompletionMethod = 0 	-
	Upon upgrade the following IPoIB Registry Keys will be deleted: <ul style="list-style-type: none"> • UseRSSForRawIP • UseRSSForUDP 	-
	The empty directory %ProgramFiles%\Mellanox\MLNX_VPI\Endure DLLs is being created during the driver installation.	-
	Uninstalling the driver on Windows Server 2008 R2 with LBFO configuration results in the appearance of a pop-up window requesting to close several running applications.	Choose "Do not close applications". This action allows the uninstallation of the driver. A Reboot may be required. Rebooting the server before uninstalling the driver when LBFO is configured will eliminate this pop-up completely.
	Upgrading the driver while perfmon is open and monitoring the "ConnectX NIC device" may end up with adapters having yellow bang with code 10.	Close the perfmon before driver upgrade.
Utilities	ibdump may encounter packet drops upon a burst of more than 4096 (or 2^max-burst) packets.	-
	Packets loss is not reported by ibdump.	-
	Running ibdump on a RoCE Ethernet port may decrease the functional bandwidth due to the overhead of creating extra copy for each packet. This may lead to packet drops on the link.	Verify Ethernet flow control is enable to ensure a lossless link
	Sniffing over IB ports is currently not supported	-

6 API Changes

6.1 API Changes in WinOF Rev 4.60

The following are the API changes in WinOF Rev 4.60:

Table 8 - API Changes in WinOF Rev 4.60

Name	Description
IB_MOD_QP_CHANGE_COUNTER_INDEX	Added a new macro
struct ib_qp_mod_t	Added the field state.rtr.counter_index

6.2 API Changes in WinOF Rev 4.55

The following are the API changes in WinOF Rev 4.55:

Table 9 - API Changes in WinOF Rev 4.55

Name	Description
RDMA_TRANSPORT_RRDMAOE_1_5	Added enumerated values
RDMA_TRANSPORT_RRDMAOE_2_0	Added enumerated values
RDMA_TRANSPORT_RRDMAOE	It is an alias to: RDMA_TRANSPORT_RRDMAOE_1_5
is_rroce(), is_xroce()	Added new functions
IB_AC_SNIFFER	Added a new macro
struct ib_qp_mod_t	Added the field state.init.flags

6.3 API Changes in WinOF Rev 4.40

The following are the API changes in WinOF Rev 4.40:

Table 10 - API Changes in WinOF Rev 4.40

Name	Description
VERBS_MINOR_VER	Increased its value, 0009 -> 000a
enum eth_link_speeds	Added enumerated values
struct ib_port_attr_t	<ul style="list-style-type: none"> The mtu field was separated into two fields: <ul style="list-style-type: none"> max_mtu (maximum MTU supported by the port) active_mtu (actual MTU which the port is configured with) Added the eth_link_speed field
WR_SEND_INV	Added enumerated values
struct ib_send_wr_t	The type of invalidate_rkey was changed from net32_t -> ib_net32_t

Table 10 - API Changes in WinOF Rev 4.40

Name	Description
IB_SEND_OPT_SKIP_DOORBELL	Added the send Write flag